Wavelets and Applications

Dan A. Simovici

University of Massachusetts at Boston,

Department of Computer Science,

Boston, Massachusetts 02125, USA

Content

1. The Haar Transform of Time Series
2. Wavelets
3. Data Compression
4. Applications
    - Relational Databases
    - Data Streams
    - Other Applications

The Haar Transform

Values of an analog signal measured at time values $1, \ldots, n$ are $x_1, \ldots, x_n$.
The *support* of the sequence **x** $= (x_1, \ldots, x_n)$ is the set $\{i | x_i \neq 0\}$.
We form two sequences of size $n/2$:

$$t_1, \ldots, t_{n/2}, \text{ and } f_1, \ldots, f_{n/2}$$

$$t_m = \frac{x_{2m-1} + x_{2m}}{\sqrt{2}} \text{ and } f_m = \frac{x_{2m-1} - x_{2m}}{\sqrt{2}}$$

for $1 \leq m \leq \frac{n}{2}$.

Example:

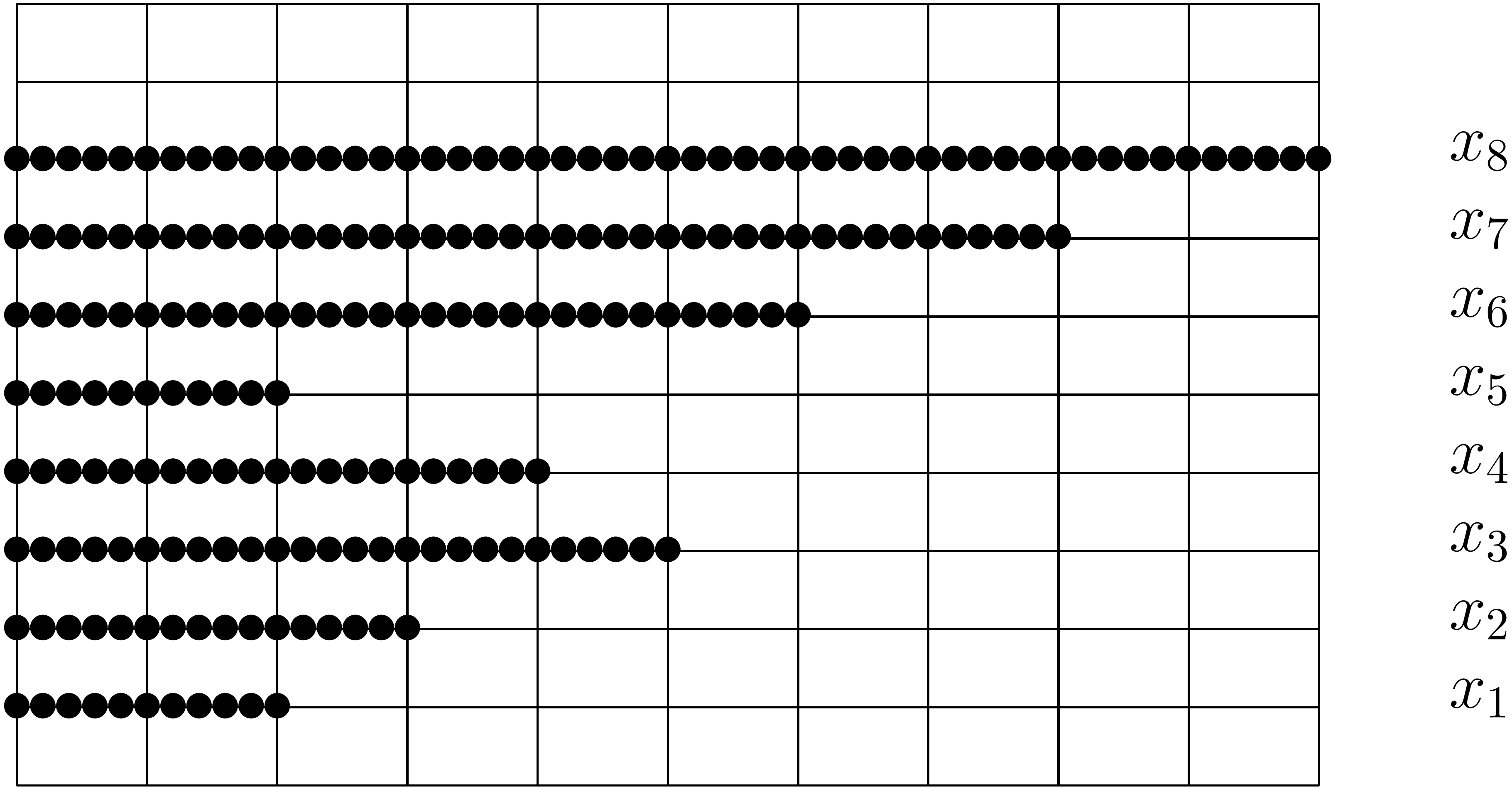
Let

$$\mathbf{x} = (x_1, \ldots, x_8) = (2, 3, 5, 4, 2, 6, 8, 10)$$

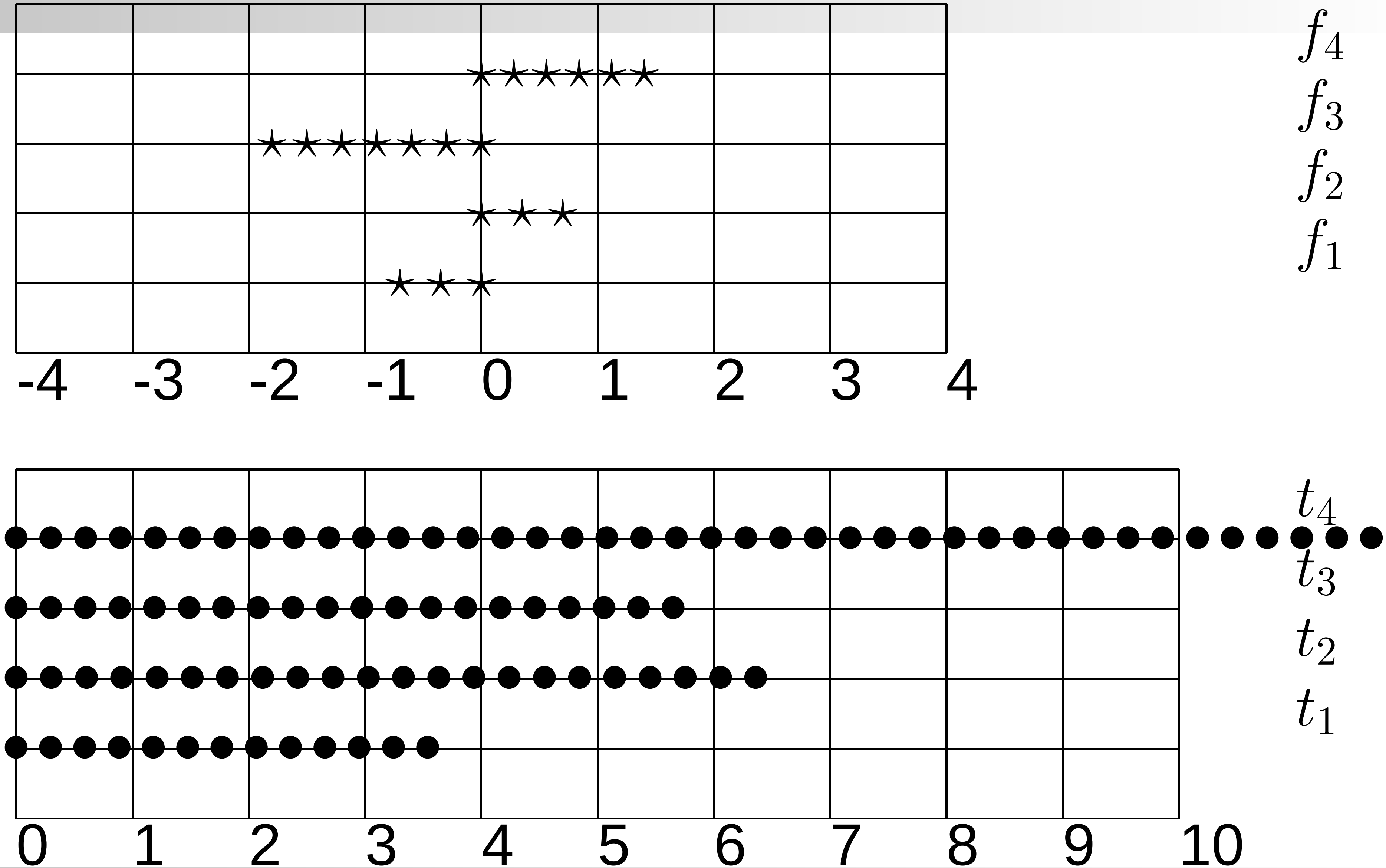
$\mathbf{t} = (t_1, \ldots, t_4)$ and $\mathbf{f} = (f_1, \ldots, f_4)$:

$$\begin{aligned} t_1 &= \frac{5}{\sqrt{2}} = 3.54 & f_1 &= \frac{-1}{\sqrt{2}} = -0.70 \\ t_2 &= \frac{9}{\sqrt{2}} = 6.36 & f_2 &= \frac{1}{\sqrt{2}} = 0.70 \\ t_3 &= \frac{8}{\sqrt{2}} = 5.65 & f_3 &= \frac{-4}{\sqrt{2}} = -2.80 \\ t_4 &= \frac{18}{\sqrt{2}} = 12.85 & f_4 &= \frac{-2}{\sqrt{2}} = -1.40 \end{aligned}$$

Original Sequence $x_1, \ldots, x_8$



Fluctuations and Trends



Remarks...

- The trend components $t_1, \ldots, t_4$ approximate the trends in **x**.
- The fluctuation components $f_1, \ldots, f_4$ approximate the fluctuations of **x**.
- Conservation of energy:

$$\mathcal{E}(\mathbf{x}) = \sum_{i=1}^{8} x_i^2 = \sum_{i=1}^{4} t_i^2 + \sum_{i=1}^{4} f_i^2,$$

- Fluctuations are small because **x** originates typically in sampling of a continuous signal.

Haar Transform

The Haar transform is the mapping $\mathcal{H}$: Seq$(\mathbb{R}) \longrightarrow$ Seq$(\mathbb{R})$ given by:

$$\mathcal{H}(x_1, \ldots, x_n) = (t_1, \ldots, t_{\frac{n}{2}}, f_1, \ldots, f_{\frac{n}{2}})$$

for $(x_1, \ldots, x_n) \in$ Seq$(\mathbb{R})$.
We use the condensed notation $\mathcal{H}(\mathbf{x}) = (\mathbf{t}^1 | \mathbf{f}^1)$, where

$$\begin{aligned} \mathbf{t}^1 &= (t_1, \ldots, t_{\frac{n}{2}}) \\ \mathbf{f}^1 &= (f_1, \ldots, f_{\frac{n}{2}}) \end{aligned}$$

The Inverse Haar Transform

If $\mathcal{H}(x_1, \ldots, x_n) = (t_1, \ldots, t_{\frac{n}{2}}, f_1, \ldots, f_{\frac{n}{2}})$, then

$$
\begin{aligned}
t_1 &= \tfrac{x_1+x_2}{\sqrt{2}} & f_1 &= \tfrac{x_1-x_2}{\sqrt{2}} \\
&\vdots & &\vdots \\
t_{\frac{n}{2}} &= \tfrac{x_{n-1}+x_n}{\sqrt{2}} & f_{\frac{n}{2}} &= \tfrac{x_{n-1}-x_n}{\sqrt{2}}
\end{aligned}
$$

Then:

$$
\begin{aligned}
x_1 &= \tfrac{t_1+f_1}{\sqrt{2}} & x_2 &= \tfrac{t_1-f_1}{\sqrt{2}} \\
&\vdots & &\vdots \\
x_{n-1} &= \tfrac{t_{\frac{n}{2}}+f_{\frac{n}{2}}}{\sqrt{2}} & x_n &= \tfrac{t_{\frac{n}{2}}-f_{\frac{n}{2}}}{\sqrt{2}}
\end{aligned}
$$

The Inverse Haar Transform (cont)

The inverse Haar transform is the mapping $\mathcal{H}^{-1}$: Seq$(\mathbb{R}) \longrightarrow$ Seq$(\mathbb{R})$ given by:

$$\mathcal{H}^{-1}(t_1, \ldots, t_{\frac{n}{2}}, f_1, \ldots, f_{\frac{n}{2}}) = (x_1, \ldots, x_n)$$

for $(t_1, \ldots, t_{\frac{n}{2}}, f_1, \ldots, f_{\frac{n}{2}}) \in$ Seq$(\mathbb{R})$.

Higher-Level Haar Transforms

For $\mathbf{x} \in \mathbb{R}^n$ and $k = \log_2 n$:

$$
\begin{aligned}
\mathcal{H}^{[1]}(\mathbf{x}) &= \mathcal{H}(\mathbf{x}) = (\mathbf{t}^1|\mathbf{f}^1), \\
\mathcal{H}^{[2]}(\mathbf{x}) &= (\mathcal{H}(\mathbf{t}^1)|\mathbf{f}^1) = (\mathbf{t}^2|\mathbf{f}^2|\mathbf{f}^1), \\
\mathcal{H}^{[3]}(\mathbf{x}) &= (\mathcal{H}(\mathbf{t}^2)|\mathbf{f}^2|\mathbf{f}^1) = (\mathbf{t}^3|\mathbf{f}^3|\mathbf{f}^2|\mathbf{f}^1), \\
&\vdots \\
\mathcal{H}^{[k]}(\mathbf{x}) &= (\mathbf{t}^k|\mathbf{f}^k|\mathbf{f}^{k-1}|\cdots|\mathbf{f}^1)
\end{aligned}
$$

The Full Haar Transform

The *full Haar transform* of a sequence **x** of length n is

$$\mathbf{H}(\mathbf{x}) = (\mathbf{t}^k|\mathbf{f}^k|\mathbf{f}^{k-1}|\cdots|\mathbf{f}^1),$$

where $k = \log_2 n$.

Energy Localization Property

Most of energy is concentrated in the trend vector. For

$$\mathbf{x} = (2, 3, 5, 4, 2, 6, 8, 10)$$

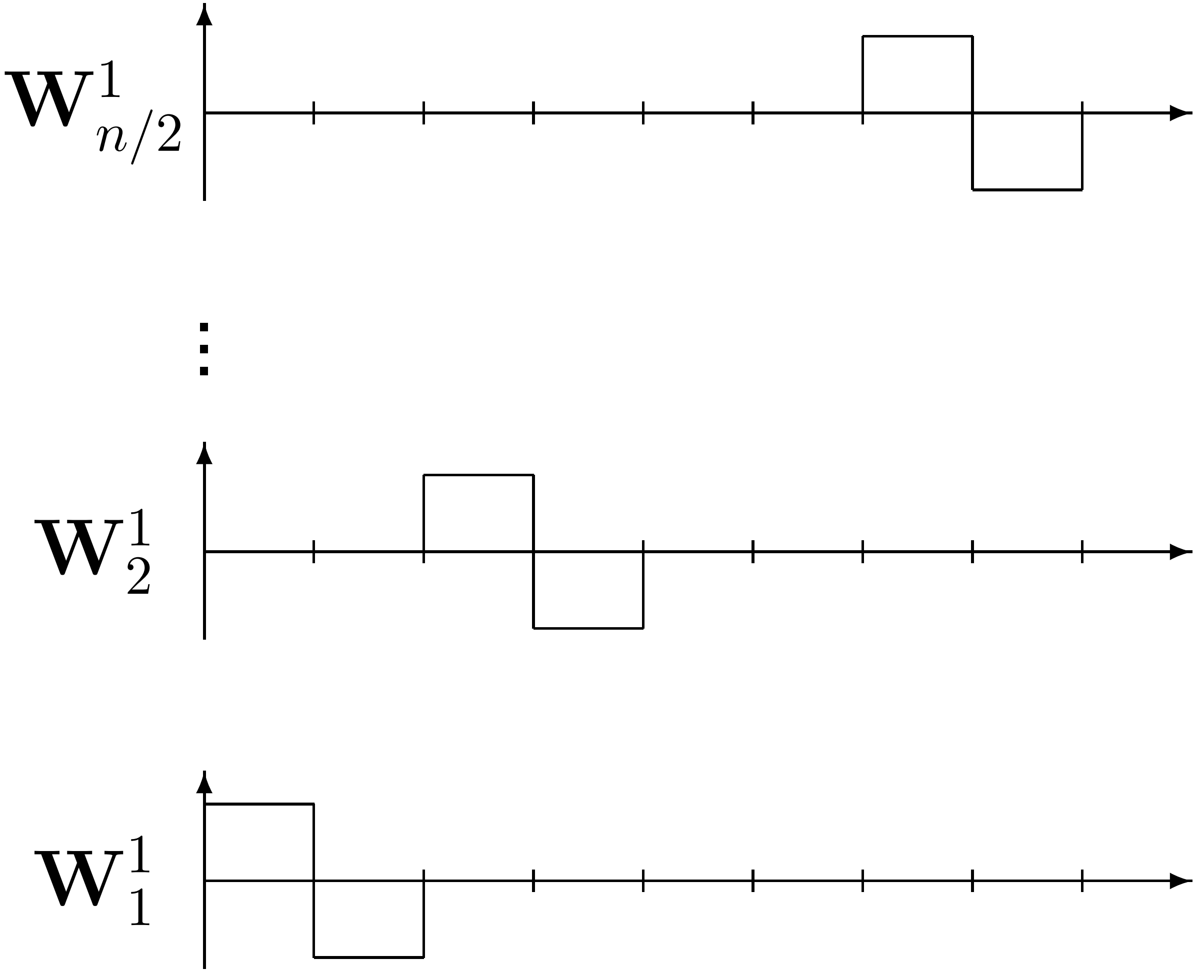
we have

$$\mathcal{E}(\mathbf{x}) = 2^2 + 3^2 + 5^2 + 4^2 + 2^2 + 6^2 + 8^2 + 10^2 = 258$$
$$\mathcal{E}(\mathbf{t}^1) = 3.54^2 + 6.36^2 + 5.65^2 + 12.85^2 \approx 246$$
$$\mathcal{E}(\mathbf{f}^1) = 0.70^2 + 0.70^2 + 2.80^2 + 1.40^2 \approx 12$$

Haar Wavelets

The *1-level Haar wavelets* are the sequences

$$
\begin{aligned}
\mathbf{W}^1_1 &= (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0, 0, \ldots, 0, 0) \\
\mathbf{W}^1_2 &= (0, 0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, \ldots, 0, 0) \\
&\vdots \\
\mathbf{W}^1_{\frac{n}{2}} &= (0, 0, 0, 0, \ldots, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})
\end{aligned}
$$


$\mathbf{W}^1_{n/2}$
⋮
$\mathbf{W}^1_2$
$\mathbf{W}^1_1$

A Bit of History ...

- Weierstrass (1873): a family of functions constructed by superimposing scaled copies of a given base function.
- Haar (1909): introduced the Haar basis (compact support).
- Gabor (1946): nonorthogonal basis of functions with unbounded support (translations of Gaussians).
- Ricker (1940): the term wavelet (seismology)

Properties of Wavelets

If $\mathcal{H}(\mathbf{x}) = (\mathbf{t}|f_1, \ldots, f_{\frac{n}{2}})$, then

$$f_i = \mathbf{x}\mathbf{W}_i \text{ for } 1 \leq i \leq \frac{n}{2}.$$

- Average value of a wavelet is 0.
- For each wavelet $\mathbf{W}_i^1$ we have $\mathcal{E}(\mathbf{W}_i^1) = 1$.
- Each wavelet can be obtained from the first wavelet by a time-translation of 2.
- If $\mathbf{x}$ is approximatively constant on $\mathrm{supp}(W_i^1)$, then f_i is approximatively 0.

The Haar Scaling signals

The *1-level Haar scaling signals* are:

$$
\begin{aligned}
\mathbf{V}_1^1 &= (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0, \ldots, 0, 0) \\
\mathbf{V}_2^1 &= (0, 0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \ldots, 0, 0) \\
&\vdots \\
\mathbf{V}_{\frac{n}{2}}^1 &= (0, 0, 0, 0, \ldots, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})
\end{aligned}
$$

Properties of Scaling Signals

If $\mathcal{H}(\mathbf{x}) = (t_1, \ldots, t_{\frac{n}{2}} | \mathbf{f})$, then

$$t_i = \mathbf{x}\mathbf{V}_i^1 \text{ for } 1 \leq i \leq \frac{n}{2}.$$

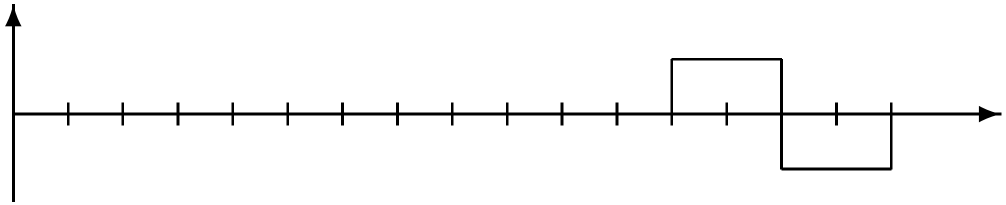
- Average value of a scaling signal is not 0.
- For each scaling signal $\mathbf{V}_i^1$ we have $\mathcal{E}(\mathbf{V}_i^1) = 1$.
- Each scaling signal can be obtained from the first scaling signal by a time-translation of 2.

2nd-Level Wavelets

The *2nd-level wavelets* are defined by

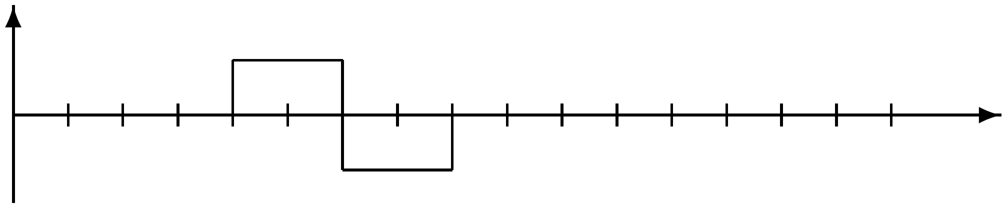
$$\begin{aligned}
\mathbf{W}_1^2 &= (\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, 0, \ldots, 0) \\
\mathbf{W}_2^2 &= (0, 0, 0, 0\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, 0, \ldots, 0) \\
&\vdots \\
\mathbf{W}_{\frac{n}{4}}^2 &= (0, 0, 0, 0, 0, \ldots, 0, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2})
\end{aligned}$$

$\mathbf{W}^2_{n/4}$

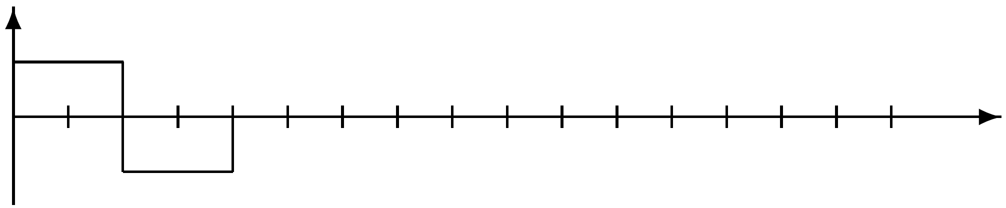


$\vdots$

$\mathbf{W}^2_2$



$\mathbf{W}^2_1$



2nd-Level Scaling Signals

The *2nd-level scaling* are defined by

$$
\begin{aligned}
\mathbf{V}_1^2 &= (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, 0, \ldots, 0) \\
\mathbf{V}_2^2 &= (0, 0, 0, 0, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, 0, \ldots, 0) \\
&\vdots \\
\mathbf{V}_{\frac{n}{4}}^2 &= (0, 0, 0, 0, 0, \ldots, 0, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})
\end{aligned}
$$

2nd-order fluctuations:

$$f_i^2 = \textbf{xW}_i^2 \text{ for } 1 \leq i \leq \frac{n}{4}$$

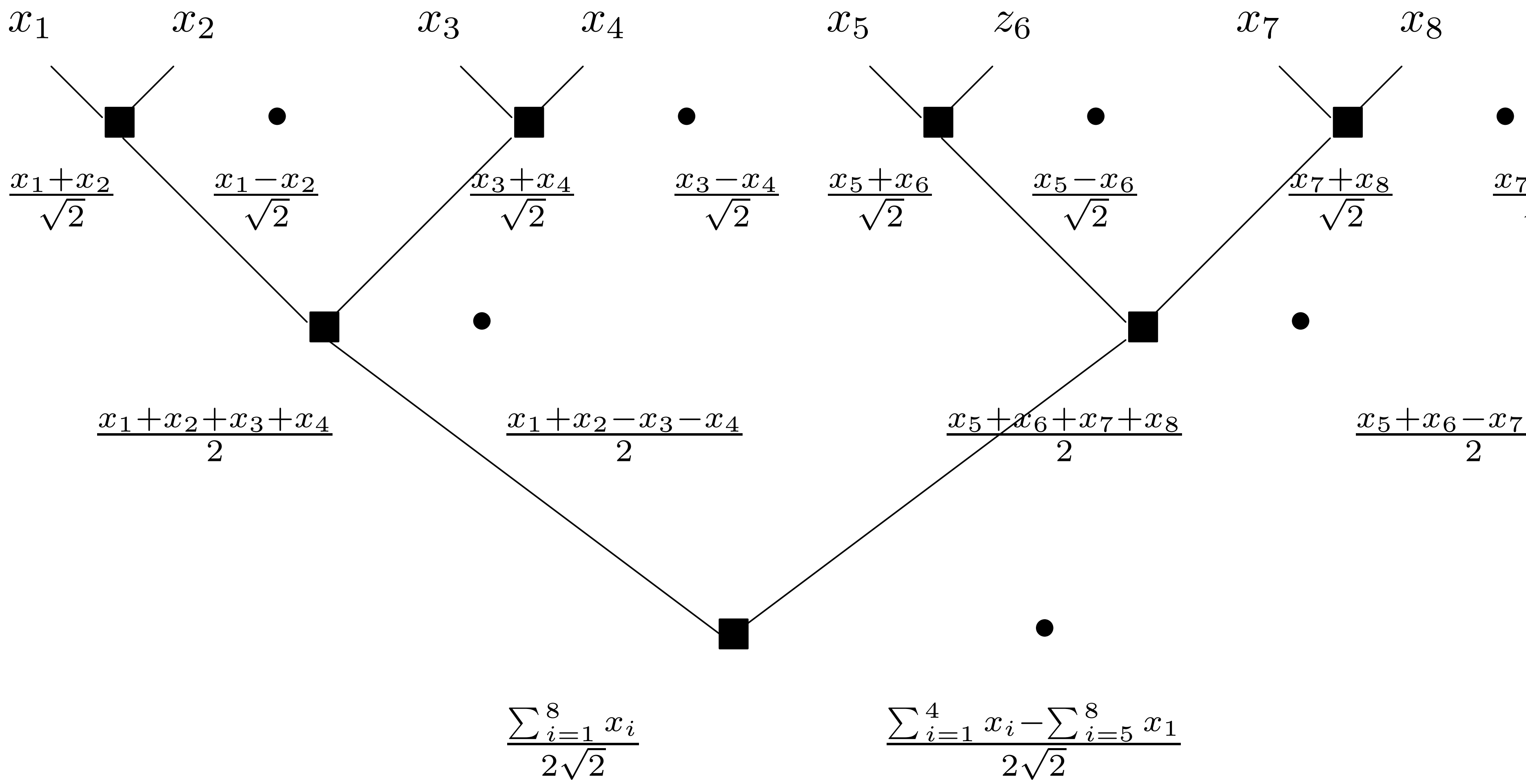
2nd-order trends:

$$t_i^2 = \textbf{xV}_i^2 \text{ for } 1 \leq i \leq \frac{n}{4}$$

Properties of the 2nd-level wavelets anc

- Average value of wavelets is 0; average value of scaling signals is non-zero.
- $\mathcal{E}(\mathbf{W}_i^2) = \mathcal{E}(\mathbf{V}_i^2) = 1$,
- $\text{supp}(\mathbf{W}_i^2) = \text{supp}(\mathbf{V}_i^2) = 4,$ for $1 \leq i \leq \frac{n}{4}$.

Computation Tree



Full System of Wavelets

For $1 \leq j \leq \log_2 n$ and $1 \leq h \leq \frac{n}{2^j}$ define:

$$\mathbf{W}_h^j = (0, \ldots, 0, \underbrace{\left(\frac{1}{\sqrt{2}}\right)^j, \ldots, \left(\frac{1}{\sqrt{2}}\right)^j}_{2^{j-1}}, \underbrace{-\left(\frac{1}{\sqrt{2}}\right)^j, \ldots, -\left(\frac{1}{\sqrt{2}}\right)^j}_{2^{j-1}}, 0, \ldots, 0)$$

Full System of Wavelets for $n = 8$

$$W_1^1 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0, 0, 0, 0, 0, 0\right) \qquad W_2^1 = \left(0, 0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0, 0, 0, 0\right)$$

$$W_3^1 = \left(0, 0, 0, 0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0, 0\right) \qquad W_4^1 = \left(0, 0, 0, 0, 0, 0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$$

$$W_1^2 = \left(\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}, 0, 0, 0, 0\right) \qquad W_2^2 = \left(0, 0, 0, 0, \tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}\right)$$

$$W_1^3 = \left(\left(\frac{1}{\sqrt{2}}\right)^3, \left(\frac{1}{\sqrt{2}}\right)^3, \left(\frac{1}{\sqrt{2}}\right)^3, \left(\frac{1}{\sqrt{2}}\right)^3, -\left(\frac{1}{\sqrt{2}}\right)^3, -\left(\frac{1}{\sqrt{2}}\right)^3, -\left(\frac{1}{\sqrt{2}}\right)^3, -\left(\frac{1}{\sqrt{2}}\right)^3\right)$$

Full System of Wavelets for $n = 8$

Full Haar transform of a sequence **x**:

$$(\mathbf{x}\mathbf{V}_1^3, \mathbf{x}W_1^3, \mathbf{x}W_1^2, \mathbf{x}W_2^2,$$
$$\mathbf{x}W_1^1, \mathbf{x}W_2^1, \mathbf{x}W_3^1, \mathbf{x}W_4^1)$$

Example: For $\mathbf{x} = (2, 3, 5, 4, 2, 6, 8, 10)$:

$$\begin{aligned} \mathbf{H}(\mathbf{x}) &= (14.2, -4.30, -1.99, -5.09, \\ &\quad -0.70, 0.70, -2.80, -1.40) \end{aligned}$$

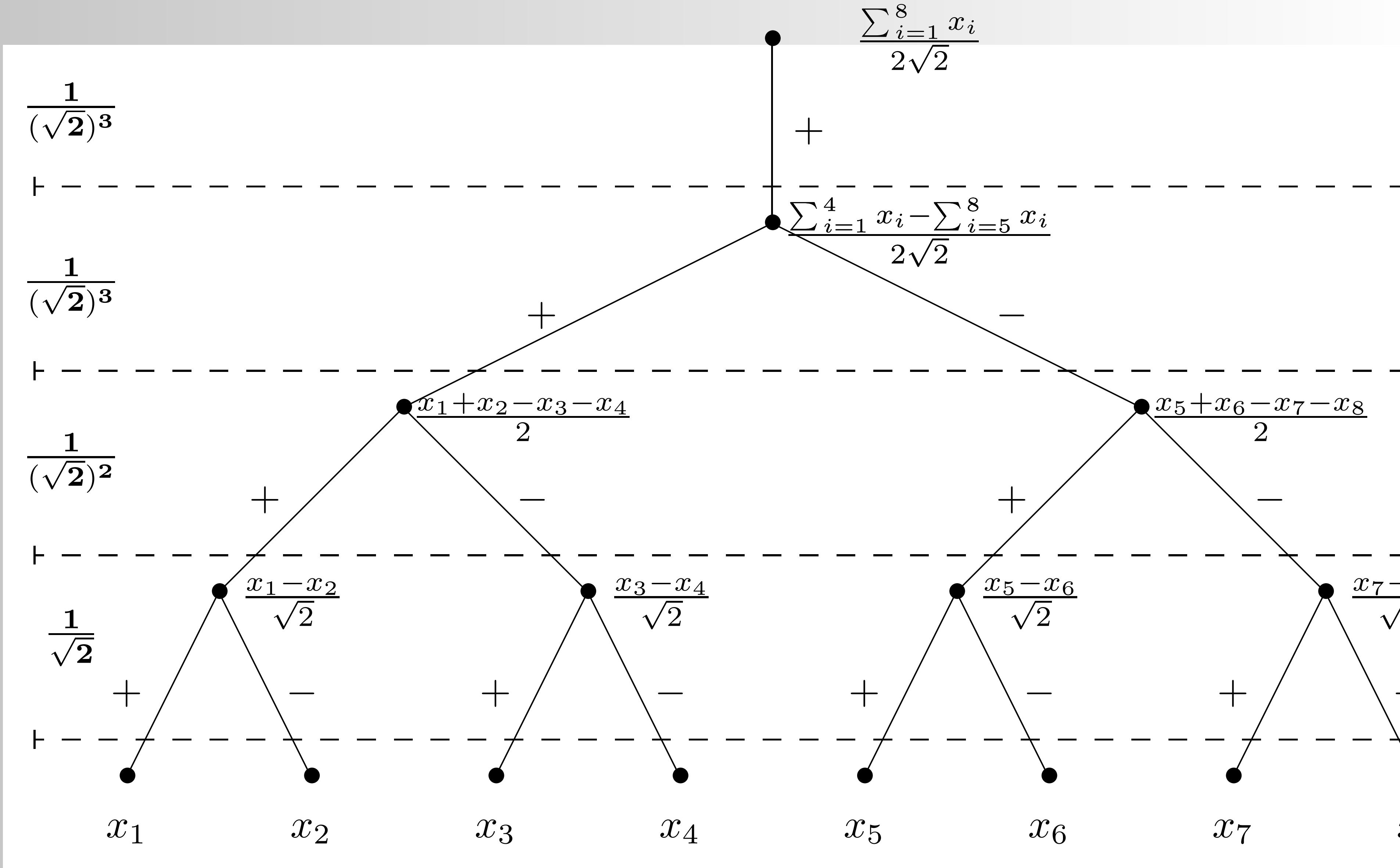
Synthesis of Signals

The inverse Haar transform:

$$x_1 = \frac{t_1 + f_1}{\sqrt{2}}, x_2 = \frac{t_1 - f_1}{\sqrt{2}}, \ldots, x_n = \frac{t_{\frac{n}{2}} - f_{\frac{n}{2}}}{\sqrt{2}}$$

$$\mathbf{x} = \left(\frac{t_1}{\sqrt{2}}, \frac{t_1}{\sqrt{2}}, \ldots, \frac{t_{\frac{n}{2}}}{\sqrt{2}}, \frac{t_{\frac{n}{2}}}{\sqrt{2}}\right) + \left(\frac{f_1}{\sqrt{2}}, -\frac{f_1}{\sqrt{2}}, \ldots, \frac{f_{\frac{n}{2}}}{\sqrt{2}}, -\frac{f_{\frac{n}{2}}}{\sqrt{2}}\right)$$

The Inverse of the full Haar Transform



Multiresolution Analysis

First averaged and first detail signals are:

$$\mathbf{T}^1 = \left(\frac{t_1}{\sqrt{2}}, \frac{t_1}{\sqrt{2}}, \ldots, \frac{t_{\frac{n}{2}}}{\sqrt{2}}, \frac{t_{\frac{n}{2}}}{\sqrt{2}}\right)$$

$$\mathbf{F}^1 = \left(\frac{f_1}{\sqrt{2}}, -\frac{f_1}{\sqrt{2}}, \ldots, \frac{f_{\frac{n}{2}}}{\sqrt{2}}, -\frac{f_{\frac{n}{2}}}{\sqrt{2}}\right)$$

$\mathbf{x} = \mathbf{T}^1 + \mathbf{F}^1$: sum of a lower resolution signal and a detail signal.

Averaged and Detail Signals (cont)

The averaged and detail signals can be written as

$$\begin{aligned} \mathbf{T}^1 &= t_1 \mathbf{V}_1^1 + \cdots + t_{\frac{n}{2}} \mathbf{V}_{\frac{n}{2}}^1 \\ &= (\mathbf{x}\mathbf{V}_1^1)\mathbf{V}_1^1 + \cdots + (\mathbf{x}\mathbf{V}_{\frac{n}{2}}^1)\mathbf{V}_{\frac{n}{2}}^1 \\ \mathbf{F}^1 &= f_1 \mathbf{W}_1^1 + \cdots + f_{\frac{n}{2}} \mathbf{W}_{\frac{n}{2}}^1 \\ &= (\mathbf{x}\mathbf{W}_1^1)\mathbf{W}_1^1 + \cdots + (\mathbf{x}\mathbf{W}_{\frac{n}{2}}^1)\mathbf{W}_{\frac{n}{2}}^1 . \end{aligned}$$

Example:

Let $\mathbf{x} = (2,3,5,4,2,6,8,10)$. We have

$$\begin{aligned}\mathcal{H}(\mathbf{x}) \;&=\; (\frac{5}{\sqrt{2}}, \frac{9}{\sqrt{2}}, \frac{8}{\sqrt{2}}, \frac{18}{\sqrt{2}}| \\ &\quad -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -\frac{4}{\sqrt{2}}, -\frac{2}{\sqrt{2}})\end{aligned}$$

The averaged signal:
$\mathbf{T}^1 = (5/2, 5/2, 9/2, 9/2, 4, 4, 9, 9)$
The detail signal:

$\mathbf{F}^1 = (-1/2, 1/2, 1/2, -1/2, -2, 2, -1, 1)$

Multiple-level MRA

$$\begin{aligned} \mathbf{x} &= \mathbf{T}^1 + \mathbf{F}^1 \\ \mathbf{T}^1 &= \mathbf{T}^2 + \mathbf{F}^2 \\ &\vdots \\ \mathbf{T}^{k-1} &= \mathbf{T}^k + \mathbf{F}^k, \end{aligned}$$

so

$$\mathbf{x} = \mathbf{T}^k + \mathbf{F}^k + \mathbf{F}^{k-1} + \cdots + \mathbf{F}^1$$

Compression of Signals

- Compression: converting a signal into a new format that requires fewer bits to transmit
- Categories of compression
  - *lossless copmpression*: error-free decompression of the original signal (Huffman compression, LZW compression, arithmetic compression)
  - *lossy compression*: produces inaccuracies in the decompressed signal
- rates of compression (50:1–100:1) for lossy compr. vs. 2:1 for lossless

Wavelet Compression Methods

1. Compute a wavelet transform of a signal.
2. Set to 0 all values of components that are below a threshold value λ.
3. Transmit only the significant, non-zero values.
4. Compute the reverse transform at the receiving end, using zero values forthe components that were not transmitted.

A Relational Database Application

Selectivity Estimation

ORDERS

cust_no	cust_name	date	qty
123	John Doe	2/10/2003	8
⋮	⋮	⋮	⋮

Find the fraction of ORDERS returned by:

select cust_name **from** ORDERS
where 1 <= qty **and** qty <= 3;

Wavelet-based Histograms (Vitter)

The active domain of A: $v_1 < v_2 < \cdots < v_n$: the values that appear under an attribute A of a table.

Frequencies: $f_i = |\{t|t[A] = v_i\}|$, $1 \leq i =\leq n - 1$

Cumulative Frequencies:

$$c_i = |\{t|t[A] <= v_i\}| = \sum_{k=1}^{i} f_k,$$

for $1 \leq i =\leq n - 1$

Cumulative Data Distribution

Data distribution of A:

$$\mathcal{T}(A) = \{(v_1, f_1), \ldots, (v_n, f_n)\}$$

Cumulative data distribution of A:

$$\mathcal{T}^C(A) = \{(v_1, f_1), \ldots, (v_n, f_n)\}$$

Extended cumulative data distribution $\mathcal{T}^{C+}(A)$ is the extension of $\mathcal{T}^C$ obtained by assigning 0 frequencies to all values that do not occur in the table.

Vitters' Histogram Construction

1. form the extended cummulative distribution $\mathcal{T}^{C+}(A)$ (preprocessing);
2. compute $\mathcal{H}(\mathcal{T}^{C+}(A))$;
3. retain only the m most significant wavelet coefficients for some m that corresponds to the desired storage usage.

The number of tuples $T(A)_{a,b}$ such that $a \leq A \leq b$ is

$$T(A)_{a,b} = \mathcal{T}^{C+}(A)_b - \mathcal{T}^{C+}(A)_{a-1}$$

Example:

ORDERS	
...	qty
	1
	3
	4
	3
	1
	4
	3
	3
	3

$$\mathcal{T}(\text{qty}) = \{(1,2),(3,5),(4,2)\}$$

$$\mathcal{T}^{C+}(\text{qty}) = \{(1,2),(2,2),(3,7),(4,9)\}$$

$$\mathcal{H}(\mathcal{T}^{C+}(2,2,7,9)) = (9.99,-5.99,0,-1.91)$$

$$\mathcal{H}^{-1}(\mathcal{T}^{C+}(9.99,-5.99,0,-1.91)) = (1.99,1.99,6.99,8.99)$$

Further steps and remarks ...

- The value of the m coefficients together with their positions serve as histogram.
- To estimate the value of $|\{t|c \geq t[A] \geq d\}|$ we construct the values for b and $a-1$ in in the extended cumulative distribution function and then take their difference.
- Effectiveness is increased when we replace the raw frequencies with $\mathcal{T}^{C+}(A)$.

Preprocessing

- If the active domain V is small, an one-pass, in-memory computation is sufficient.
- If V is large, use an external merge-sort and sum up the frequencies of the records that are merged.
- If V is very large use random sampling and use the sample data distribution as an approximation.

Restricting the Coefficients

Thresholding: m out of N coefficients are kept; the remaining are set to 0. Then, the inverse Haar transform is computed.
Let s be the size of query q and s' be the size of query q after thresholding.
Error computations for a query q_i:

- absolute error: $e^{abs}(q) = |s - s'|$ (small freqs.)
- relative error: $e^{rel}(q) = \frac{e^{abs}}{s}$ (large freqs.)
- combined error: $e^{comb}(q) = \min\{\alpha e^{abs}(q), \beta e^{rel}(q)\}$

Global error for a set of queries

For a set $Q = \{q_1, \ldots, q_k\}$ of queries we have an error vector

$$\mathbf{e} = (e(q_1), \ldots, e(q_k))$$

The overall error is

$$||e||_p = \left(\frac{1}{k} \sum_{i=1}^{k} e_i^p \right)^{\frac{1}{p}}$$

Thresholding Techniques

- Choose the largest m wavelet coefficients in absolute value.
- Choose m coefficients in a greedy way (e.g. as above), then repeatedly include the coefficients that decrease the error and exclude those that increase it.

Estimating Selectivity

Vitter's Theorem: For a given range query $a \leq X \leq b$, the cumulative frequencies of $a-1$ and b can be reconstructed from m wavelet coefficients using $O(m)$ space in time $O(\min\{m, \log N\})$.

Mining Data Streams

Mining data that arrives and is processed in a stream: *"you look only once"*
Examples:

- switches and routers in networks generate data on
  - telephone calls
  - IP addresses
- streams of credit card transactions
- log records in web-based services

Main Challenge:

Data accumulation is expensive so it is important to extract information even at the cost of obtaining approximative results.

The Processing Model

Characteristics of stream processing are identified:

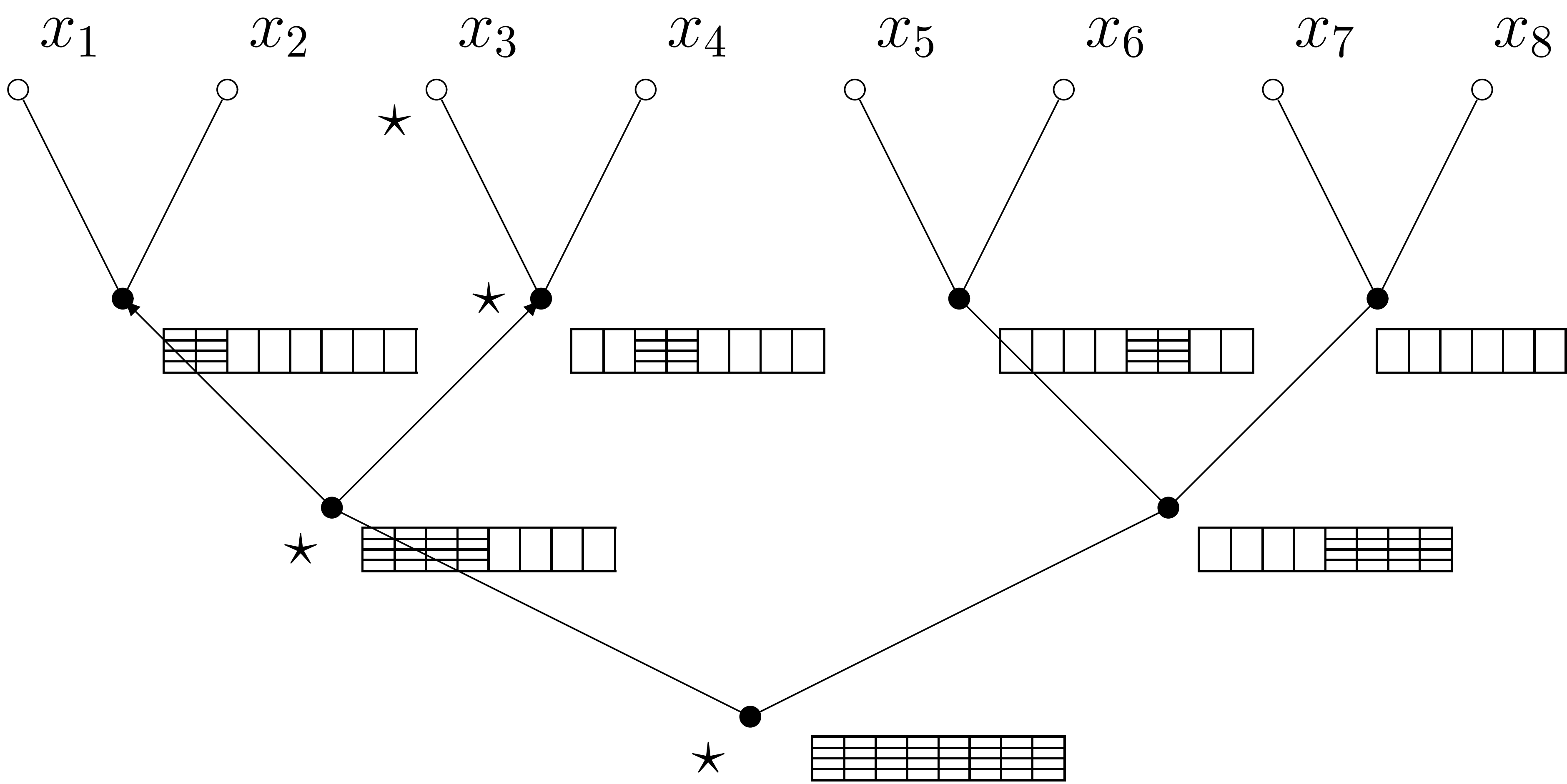
- each data item is read and processed as soon as it arrives;
- no backtracking is allowed on the data stream;
- explicit access to arbitrary past items is not allowed.

What is allowed ...

An additional amount of memory is permitted subjected to the following conditions:

- the additional memory may be used to store:
  - a recent window of items;
  - some sumary information about the stream.
- the size of the memory is significantly smaller than the signal domain size.

Straddling Coeficients



Computation of the higest m terms

The highest m terms yields the best approximation for the error $||e||_2$.

Gilbert's result: With the most $O(m + \log N)$ storage we can compute the highest m-term approximation to a signal. Each new data signal item needs $O(m + \log n)$ time to be processed.

Lower Space Bound

Any streaming algorithm that correctly calculates the highest wavelet basis coefficient of a signal requires $\Omega(\frac{N}{\log \log N})$ space.

Other Applications

- Clustering time series that represent levels of gene expressions in microarrays as they appear in the mitosis process (a study of cellular division of the cells that form the retina).
- The new image data compression standard JPEG 2000

Conclusions

- Wavelet transforms generate simple algorithms for data compression.
- Computations can be done efficiently, in small space.
- A large variety of applications exist even for the simplest Haar wavelets.